BREVIORA

Museum of Comparative Zoology

US ISSN 0006-9698

JAN 23 1995

CAMBRIDGE, MASS.

10 January 1995

Number 502

A COMPUTER APPROACH TO THE COMPARISON AND IDENTIFICATION OF SPECIES IN DIFFICULT TAXONOMIC GROUPS*

ERNEST E. WILLIAMS, 1,5 HUGH RAND, 2 A. STANLEY RAND, 3 AND ROBERT J. O'HARA4

ABSTRACT. A computer program for the identification of unknown taxa in "difficult groups" based on matching rather than sequential exclusion is proposed as a substitute for both the conventional dichotomous key and for the random entry matrix that has been suggested as a replacement for the dichotomous key. The matching program is modeled after the steps that a practicing taxonomist would employ in the identification of an unknown specimen:

- 1. Data for the unknown are compiled.
- 2. Data for the unknown are compared with those of relevant known taxa.
- 3. On the basis of the comparison, certain of the named taxa are considered possible matches with the unknown.
- 4. Final choice of the named taxon best matching the unknown is made and confirmed from additional data. If there is no match, the possibility of an undescribed species must be confronted.

^{*} Apple, HyperCard, and Macintosh are registered trademarks of Apple Computer, Inc. Turbo BASIC is a registered trademark of Borland International, Inc. IBM is a registered trademark of International Business Machines Corporation. MSDOS is a registered trademark of Microsoft Corporation.

¹ Museum of Comparative Zoology, Harvard University, Cambridge, Massachusetts 02138.

² Department of Applied Mathematics, University of Washington, Seattle, Washington 98195.

³ Smithsonian Tropical Research Institute, P.O. Box 2072, Balboa, Republic of

⁴ Center for Critical Inquiry in the Liberal Arts, University of North Carolina at Greensboro, Greensboro, North Carolina 21412-5001.

⁵ To whom reprint requests should be addressed.

The program enables a computer to follow these steps:

- 1. It first enters the states of selected characters for the unknown individual or series. For each character of the unknown individual or series, two numbers are recorded, a minimum and a maximum; in the case of the series allowing for variation within the sample, and in the case of the individual allowing for differences between its two sides *and* also obscurities and ambiguities in counts or coding.
- 2. Once these data are entered, the program compares them against the ranges recorded in a reference matrix for selected known species. The user specifies the maximum number of characters in which a taxon in the reference matrix is allowed to differ from the unknown before being accepted as a "match."
- 3. The resulting report lists all the matching taxa and, for each, the number of characters not matched, the specific characters not matched, and by how much.
- 4. Included in the report for each matching taxon is a "descriptor" that cites characters that are not coded for the computer as well as characters regarded as "diagnostic" of the taxon. The descriptors assist in the final choice of the most plausible identification for the unknown. In certain cases (e.g., a new taxon), evaluation of the descriptors may require the user to reject all matches.

While the program was inspired by problems encountered during exploration of the systematics of anoline lizards, it does not deal with phylogeny at all. It is only—in our eyes—a better substitute for the dichotomous key. It aids in the identification of animals. As such, it has been customized for the anoles. The reference matrices, character descriptions, and "descriptors" provided as examples in the second and succeeding sections of this paper are for anoline lizards only. The concept of the matching program is, however, applicable to taxa of any sort.

I. HISTORICAL INTRODUCTION

ERNEST E. WILLIAMS

Some years ago, while visiting Stanley Rand in Panama, I wondered aloud whether and to what extent computers could be used in species identification. I expressed dissatisfaction with the computer keys I knew about and with the usual dichotomous keys employed in taxonomic works in general. The dichotomous key did not at all resemble the process by which the working taxonomist, engaged, for example, in revising a large and complex genus, would actually employ to distinguish the species.

In my own work on anoline lizards, I had grown into the habit of using a standard data sheet for almost every anole specimen that I encountered, whether I recognized the species on sight or not, and very definitely for any specimen that I did not recognize with complete certainty. I routinely recorded museum number and locality and then a set of about 25 characters, some of them counts, sometimes brief descriptions, and, last of all, some very brief remarks on color as preserved. I used these data sheets to compel myself to always take the same data on all specimens and to standardize my information with regard to species that I did not know well. In the case of West Indian species, which ordinarily I knew in life, color, shape, and some scale differences that seemed diagnostic were the way that I distinguished species, but with mainland species that I knew only as museum specimens—and furthermore as specimens preserved, as the older specimens usually were, in ways that minimized color distinctions, and having scale characters that I had little acquaintance with—my data sheets seemed the only way to go.

My method of comparison for these unfamiliar species and characters was, however, primitive: a random, intuitive search for matching characters and populations. It seemed to me that there must be a way to automate the comparisons—the matching—that I was already doing by hand and that computers must be the best way to do this.

Not all of this do I know to have been said at one time, or even in one place, but it is certain that at one supper hour Stan Rand agreed with my general philosophy and, with his son Hugh's concurrence, voiced the opinion that a computer matching program of the sort that I imagined was feasible. By my visit the next year, Hugh Rand had written a Basic program for the Apple II computer that was the first version of the system described here. I provided the first reference matrix, one for Puerto Rican anoles. Stanley Rand emended the Apple II version and then created an IBM PC version, for which I have added matrices for Panamanian, Costa Rican, and Ecuadorian anoles to the one for Puerto Rico. This version is used at the Smithsonian Tropical Research Institute.

Much later, when Robert O'Hara was my assistant and especially my advisor-expert for the Macintosh computer, he introduced me to HyperCard, and using that very special application he created for me the Macintosh version that is also the database on which I work and from which I furnish the reference matrices used by the IBM PC.

II. A COMPUTER MATCHING PROGRAM FOR USE AS AN IDENTIFICATION KEY

ERNEST E. WILLIAMS, HUGH RAND, AND A. STANLEY RAND

Difficult taxa are those for which none but the specialist can identify species with any confidence, and the specialist not without tremors of unease. They are typically large genera in which species distinctions are subtle or ambiguous. They may even be relatively easy to identify in one sex or under special conditions. Typically the species list is still very incomplete, and the possibility of encountering a new species is still uncomfortably high. These are taxa for which the dichotomous key may be as often a trap as an aid. It is our thought that a computer key of a novel type will be able to help where conventional dichotomous keys fail.

It was recognition of the difficulties associated with certain taxa that led Peters and Orejas Miranda (1970) and Peters and Donoso Barros (1970), in the two-volume *Catalogue of the Neotropical Squamata*, to present character matrices instead of dichotomous keys in three cases (the snake genera *Atractus* and *Bothrops* in Volume 1, and the lizard genus *Anolis* in Volume 2) and to advocate computer sorting. It is interesting to quote in full their reasons for doing so (pp. vi–vii, identically paginated in the snake and lizard volumes):

"Most of the keys presented here are the standard dichotomous type In the case of very large genera, however, we have introduced a different concept. Any attempt at writing keys for poorly known large genera is likely to be futile, and we have avoided this by presenting as much data as possible in the form of a matrix. This permits 'random entry' identification, for the user, in the matrix that he wishes to check and eliminates all taxa that do not possess that character, finally arriving at a considerably reduced number of taxa (hopefully only one) after checking a series of characters. This concept has formed the basis of computer identification, since the machine can do such sorting more rapidly and efficiently than the human, and the random entry matrices presented here are organized in such a way that they can be incorporated in the computer programs now available for such machine sorting. It is our assumption that this method

of identification will be used more and more in the future, and we hope that presenting these matrices will encourage others to begin to organize their data similarly, thus anticipating the day when computer keys are available to all users."

The matrices provided by Peters, Orejas Miranda, and Donoso Barros partly parallel the procedure we shall advocate below, but to our knowledge they have received little use. The sampling for the contained species was inadequate, a major problem for genera in which high individual variability is a characteristic. The data for our own matrices were collected from 20 specimens whenever possible and as many as are available in all other cases.

More fundamental, however, is a conceptual difference in method between the Peters, Orejas Miranda, and Donoso Barros approach and our own. Their method (and Peters advocated similar techniques in other papers; see Peters and Collette, 1968; Morse et al., 1971) was to sort by sequential exclusion: i.e., by the fundamental technique of the dichotomous key, but here pursued by "random entry" with the aid of a computer. The goal of such a method is elimination of all but one of the possible choices, and no individual of a species is allowed to have the alternative of the successively chosen character. Our technique is the inverse of this: matching, instead of exclusion. Variation and overlap are expected, and the goal is maximum congruence, rather than elimination.

We have called our system a "key." It is certainly not a conventional "key," and, although it will assist in the identification of species, it is not at all comparable to a conventional dichotomous key. Instead, it is a computer-assisted procedure deliberately modeled after the steps that a taxonomist would employ in attempting the assignment of new specimens to the recognized taxa for which no dichotomous key existed. No taxonomist reviewing a large snake genus, for example, would begin by relying on the dichotomous keys prepared by his or her predecessors. The taxonomist would begin by collecting data—data on the number and, perhaps, shape of head scales; on the number of dorsal rows and ventrals; on the pattern on the head, body, and the tail; on sexual dimorphism, locality, and habitat; etc. Thereafter, the procedure would be to associate populations of phenetically sim-

ilar animals, animals similar in detail, even if not identical, since nothing biological is exactly similar. The taxonomist would, in fact, be matching individuals and populations, determining in what regards and by how much the members of selected species might differ. Judgment would necessarily be involved. The expertise for which the experienced taxonomist is known is necessarily a familiarity gained over time with the chosen taxon. During this learning period, there grows an appreciation of which characters are most meaningful for species distinction or, at the least, the most readily determinable without major error. There comes also an appreciation of the kind and extent of the variation that seems to accord with the gaps that separate some underlying biological realities that are called "species."

The steps that the practicing taxonomist would use in the identification of a series of some unknown taxon for which there happened to be no conventional key would certainly be the following:

- 1. On the basis of prior experience, preliminary investigation, and/or literature search, characters are selected to be routinely checked in all specimens.
- 2. A table or data matrix is prepared recording the ranges of the states of the chosen characters in the named taxa to be compared.
- 3. The new specimens are examined for as many as possible of the chosen characters.
- 4. The table or data matrix is searched for matches and mismatches with the characters of the new specimens.
- 5. Tentative assignment of specimens to recognized taxa is made on the basis of closeness of match.
- 6. The tentative taxonomic assignment is confirmed or rejected with the aid of additional characters, whether from the literature or from previously determined specimens.

Our computer "key" breaks this procedure into two parts.

The *first part*, corresponding to steps 1 through 5, depends on *reference matrices* that contain codings for qualitative character states or standard counts, such as those that are (or should be) taken routinely on any specimen of the relevant taxa. These counts

or character states are stored and compared by the computer with the data entered for specimens or series to be identified.

The second part, corresponding to step 6, is a supplementary set of species descriptors (automatically provided in the computer search report) that highlights the most valuable diagnostic features of each species and reports its known geographic range, ecology, behavior, or other significant features such as color in life.

Together these two data sets are intended to permit the "keying out" of all species presently known for any taxonomic group from any studied area. Importantly the program has three especially useful capabilities:

- 1. The program will report not only those species in the reference matrices that match the specimens or series under examination but also those that are similar. The computer will, in fact, ask the user of the key by how many characters the unknown may be allowed to differ.
- 2. The program will report how many and which of the characters do not match the coded data for species in the reference matrices and by how much they differ.
- 3. The program can be used to discover shared characters—whether plesiomorphic or synapomorphic—or combinations of shared characters of either kind simply by employing as the unknown the relevant character state or states and searching for those taxa that match.

More importantly, however, our computer key contains provisions for its own improvement. Built into the program are procedures such that both the reference matrices and the descriptors may be modified or expanded to include new or more useful characters or newly discovered species or species of other regions.

When new species are discovered, a dichotomous key can be very hard to change in ways that accommodate the new data. Often old dichotomies no longer work. The addition of the new species may require that the key be rewritten entirely, and this may be a job that compels reexamination of all the taxa involved. In contrast, in our program data matrices for the new species can easily be constructed and added to the reference set with no changes to the program itself or to the other reference matrices.

It is to be emphasized that this system routinely provides for variation and ambiguity. The expectation of variation is, in fact, built into the system. All character states or counts *must* be specified by two numbers. If the character state or count is monomorphic for any species, this fact is entered in the matrix by simple repetition of the coding number (e.g., 1,1 or 3,3). If there is variation between the two sides of an animal *or* within a population or if there is *intermediacy or ambiguity* in the assessment of a character state or count, then the two appropriate extreme numbers should be entered (e.g., 1,3 for a character state or 16,19 for a count).

There are caveats that need mention. Several negative features of our "key" are intrinsic to its concept and therefore irremediable without discarding the concept. These are as follows.

1. The computer accepts as a match the minimum or the maximum recorded in the matrix or any number that falls between these extremes. This results in three problems.

First, in the three characters that involve enlarged scales (characters 7, 14, and 21), coding has to deal with conditions ranging from 0 (no scales enlarged) to a condition in which there is a series of gradually enlarging scales that must be coded arbitrarily. We have chosen to code such gradually enlarging scales as 30,30 or 50,50. This means that any species or series of an unknown where some individuals have no enlarged scales and others have many very slightly enlarged scales, or any unknown in which the choice between these two is ambiguous might be coded 0,30. Any species with this coding in the reference matrix will be matched by any unknown with an intermediate coding for this character. and any unknown with this coding will match all species in the reference matrix on this character; i.e., there will be many false matches. The descriptors should resolve this difficulty. We advise the user to treat matches that involve these enlarged scale characters with caution.

Second, many species also have overlapping ranges for individual characters but, although the means may differ greatly, the computer will, of course, accept as a match to both species any value that falls within the zone of overlap. Because of this aspect of the program, individuals will often be unambiguously identified, whereas series, even when subjectively the same species, may

not be. It must also be stressed emphatically that a report of zero difference by our program is *not* a statement of species identity. It is only a statement of zero difference in the chosen characters.

Third, in some cases zero differences appear to indicate true relationship, but many times species that are judged to be remote from one another on the characters *not* in the reference matrix are reported by our program as close or as exhibiting zero differences.

Remedies: In all cases of ambiguity on the initial run-through, recourse must be had—special consideration must be given to the descriptors, the second feature of our program, which cite such features as locality or color that are not readily computer-codable and that highlight special morphological, behavioral, or ecological traits.

Often the descriptors will eliminate taxa matched by the computer from further consideration. The excluded species may differ radically from the unknown in color or shape or in some other way not coded in the reference matrix. Alternatively, the descriptors may afford clear support for an identification by calling attention to or emphasizing very evident characters of the unknown—such as color, shape, or some feature special to the species. In any case, both tentative mismatches or matches should *always* be corroborated by examination of previously identified specimens and/or a check of the pertinent literature.

2. Instead of matching too many species, the unknown will sometimes not match any closely. If among the reference matrices there are unique types or small series, there will *ordinarily* be a failure to perfectly match a specimen or small series that, in the judgment of the inquiring systematist, should be conspecific. This is expectable. No unique biological specimen is ever exactly matched by another. In consequence, if any species is described from a single specimen, the identification of the next specimen is *always* a matter of extrapolation. A type or any unique specimen may set the universe of discourse, but it does not bound it. A type is an *example* only; it necessarily does not and cannot exhibit the range of variation that is contained in the population of which it is a member.

Remedies: In cases of failure to match any descriptor, the very real possibility of a new species will have to be confronted. Here,

as in the case of too many matches, a conservative approach is appropriate.

Special consideration should be given in both the cases of too many matches and of no match at all to unique features of the unknown, whether it is an individual or a series. Decision, in any case, depends on differences between populations. A unique type or potential type needs to be compared with populations, optimally with relevant sympatric or adjacent populations. If the characters of any unique type or potential type fall within the set that is characteristic of a well-sampled population, and the descriptors do not exclude it, synonymy is the first and most probable decision. We say only probable decision, because, before any decision, every kind of error must be ruled out. In all cases, the characters reported in the descriptors need to be carefully scrutinized—in particular color, pattern, and geography, including sympatry or near sympatry, and habitat.

3. It will never be possible to code all characters of possible taxonomic significance. Some, like color, pattern, or body shape, are difficult to code for computer use unless broken down into fine details. Every taxon will have its special features, which may often be more readily expressed verbally rather than in computer code. Some of these features may be unique and make the species instantly identifiable. (This is, in fact, a danger, because unique and easily recognizable characters may often conceal even species differences among populations so easily recognized by one character that no further scrutiny is given them.)

The flexibility of our program should be strongly emphasized. It permits the distribution of single character states or of special combinations of character states to be tracked through entire taxa or subsets of them, making possible tests of the validity of old taxa or providing evidence for the erection of new taxa. In general, the system here presented will prove readily adaptable internally to individual needs. Certainly the system itself can be modified or superseded.

Some final statements should be made. First, our program provides no escape from matters of judgment. In the characters employed as well as the species recognized, human judgment has entered in. The program can only mechanically sort and match what human judgment has given it. The program will not, in any

genuine sense, solve any taxonomic problems; it can, however, serve the purpose of calling attention to and exploring problems. One of us (EEW) now routinely uses it for the exploration of problems in the systematics of anoline lizards.

Second, it is important to emphasize that our program does not deal with phylogeny; it is a phenetic program, dealing with overall similarity. Homoplasy may often be a reason that species estimated to be remote phyletically may be reported with zero difference by our program. In such cases, rather than discovering true relationships, it discovers ecomorphs (Rand and Williams, 1969; Williams, 1972, 1983). However, it may also reveal the often subtle phyletic constraints that require differences between ecomorphs.

Finally, in the subsequent sections of this paper, the single example provided is one herpetological taxon, the anoline lizards. This is only an historical accident. The idea of a matching key arose, as recounted in Section I, among students of these peculiar and special animals. No feature of the fundamental concept of our system restricts it at all to this specific taxon. Other taxa and other environments may use our system.

For example, Knowlton (1993) has emphasized the abundance of sibling species (="difficult" genera) in marine environments. She cited this (p. 189) as due to "both inadequate study of morphological features of living organisms ('pseudo-sibling species') and divergence in habitat, life history, and chemical recognition systems without parallel divergence in morphology."

There is no reason not to extend this generalization to terrestrial systems. One of us (EEW) has been very aware in anoline lizards of "inadequate morphological study," including very inadequate and very incomplete descriptions of morphology. (The section in this paper that deals explicitly with anoline morphology is an attempt to describe some of the characters that should routinely be considered in every description of an anoline lizard. More can and should be added.) EEW has also been aware of habitat, life history, and visual (color and behavioral) recognition systems in anole lizards.

In the actual anoline matrices, only *external* morphological characters have been utilized. The ecological and color characters have been relegated to the descriptors. This is not at all a necessary

action. It may not even have been a prudent one. Any codable feature, ecological, behavioral, skeletal, or biochemical, may be utilized by our system. We sincerely hope that our matching key will be utilized by students of marine organisms as well as by students of other terrestrial biota.

III. ANOLINE LIZARDS AS A TEST CASE FOR THE MATCHING KEY

ERNEST E. WILLIAMS

The group first explored with the aid of our system is the anoline lizards, a monophyletic group that may include more than 400 species. We are able at this time to provide data for only limited segments of this chosen group. These small segments are provided only as examples and samples: examples of the method and samples simultaneously of some general problems and, of course, of the special features that must be tailored to each individual case.

The anoline lizards have inspired the development of our system because they fit so well into the category of difficult taxa. They have, indeed, traditionally been so regarded. In part—the most trivial part—this is a result of the high number of species. But, especially, the difficulty in identifying the anoles is a result of overlapping variability in scale characters and the absence or near absence of the invariant characters so necessary for the classic dichotomous key. The dearth of invariant characters is especially egregious for females and juveniles, and this by itself tends to make a dichotomous key inoperable.

In addition, the anoline lizards are highly visually oriented, and color vision is important in social interactions. Unfortunately, the importance of color in life to the animals themselves does not at all help in sorting out preserved specimens. Species often differ very little in any external feature except color, yet the usual methods of preservation obliterate or obscure characteristic colors and patterns or, alternatively, reveal patterns unusual in the live animal. The colors of the living animal are always more or less altered and may be darkened in poorly preserved specimens of many species to a uniform muddy brown. The dewlap, it is true, is usually much less altered, and especially less darkened, than

other portions of the body, but it is much more often than not quite absent in females and, unless preserved fully extended, is of little use even in males. Furthermore, even live animals may differ greatly in color or pattern according to physiological and psychological states. It is therefore small occasion for surprise that keys that rely heavily on color characters in anoline lizards are less than universally useful.

This is true, unhappily, for our matching program as much as for dichotomous keys. For species distinguished primarily on color characters, many formaldehyde-preserved specimens are useless. Our matching program may, indeed, be able to do more than the conventional dichotomous key: the 37 external characters that are used in the key may sometimes achieve identification when color, because of artifactual uniformity, fails entirely. But even scales may be poorly visible on badly preserved specimens, and it should be emphasized that old specimens of anoline lizards that are uniformly brown or black as preserved, without information on colors in life, or dry, damaged, or inadequately preserved are, for many or most purposes, worthless.

So species-rich and widely distributed a taxon as the anoline lizards is for many purposes unmanageable as a unit. Since Boulenger's key of 1885—never very useful and now, in the most genuine sense, hopelessly out of date—no dichotomous key for the anoline lizards has ever existed, nor is one ever likely to exist. It is the plausible practice of the taxonomist to break such a large unit into smaller segments, at least geographically, perhaps by political boundaries, or perhaps by presumed or provisional "natural" units. Even the computer system that we propose will work most usefully and easily on smaller units than the 400 or more species that is the expected total for the whole taxon.

As a first—but, we feel, a fair—test of our system, we have tried it on the 12 species occurring on the Puerto Rican bank: the land area exposed at various times during the Pleistocene of which Puerto Rico and the Virgin Islands and a few other small islands are the present emergent parts. Eleven of the species are unquestionably valid, ten of them—all those of mainland Puerto Rico—widely overlapping geographically or intimately interdigitating. The eleventh (A. roosevelti Grant) is rare or extinct, known only from six specimens. It was, indeed, known from only two indi-

viduals, until the recent reexamination of old collections. It is, however, distinctive enough that its species status has never been questioned. A twelfth species, *A. ernestwilliamsi*, recently described (Lazell, 1983), has a very small range, a single cay, an enclave within the range of its very widely distributed closest relative; its status is still disputed. Indeed, only the latter is a problem species in the sense that its species status *can* be questioned; all other species are somewhere sympatric with their closest relatives, *A. cuvieri* and *A. roosevelti*, which are allopatric and are distinguished by strong and consistent morphological features.

The twelve species include several species groups and range from siblings to taxa, the relationships of which, at least within Puerto Rico, are quite unclear. Small as the anole fauna of the Puerto Rican bank is, it will exemplify many of the problems of "difficult" taxa.

The anoles as a group cannot be specified by any single character nor by any unique combination of characters. The group is specified by the combination of two or three characters, any one of which may be missing. Again, it is usual, or at least not infrequent, that subgroups within the anoles present the same problem of definition. We suggest that this is the sort of difficulty that makes "difficult" taxonomic groups difficult. The special case of Puerto Rican anoles is in some regards an atypically favorable case. There are no known hybrids and no equivocal species except the one that has an extraordinarily restricted geographic range. Using both data sets, there should be in the case of the Puerto Rico anoles no specimens for which the determination is doubtful. There is negligible probability of the discovery of new species. This will not be true for other geographic segments of the anoline lizards. For these, there will be other problems to confront case by case. For example, in the case of the Panama-Costa Rica matrix, we know of at least nine Panamanian species that are undescribed and one or two others that are possibly valid; we would also add to species recognized by Savage for Costa Rica some that he has synonymized or left undescribed. The existence of undescribed or disputed species is a very general problem in all the Central American and South American areas of anole distribution. Only Cuba presents many similar issues in the West Indies.

IV. THE MORPHOLOGICAL CHARACTERS USED IN THE ANOLEKEY DESCRIBED, DEFINED, AND ILLUSTRATED

ERNEST E. WILLIAMS

The 37 counts and character descriptions provided in the matrices used by the IBM and Macintosh ANOLEKEYs have been chosen by myself as those believed, after long trial, to be the most useful examined and *readily codable* characters (size, proportions, scale character states, and counts) for anoline lizards as a whole. This is a personal judgment and perhaps a very temporary one. It can be justified only by some measure of empirical success, and it remains to be widely tested.

The following counts and character states are presented as a coded sequence of alternatives (Table 1). It will always be possible, and often probable, that counts a little higher than or states adjacent to those presented for any species in the coded matrix will be found in a specimen presented for identification. In considering whether to accept a match or a failure to match in an ANOLEKEY Report, allowance should always be made for this possibility. However, counts or conditions numerically farther away from those predicted by the matrix for the species found closest to the examined material on first trial should be regarded as evidence for preferring the identification of some other species. Certainly this is grounds for reexamining, perhaps recoding, some of the characters of the unknown anole and, as well, very carefully considering the alternative possibilities suggested by the supplementary *descriptor* provided for each matched species in the report.

It may well be that many current species matrices will require modification with an increase in sample size. (Those matrices based on samples below 20 will certainly need modification.) But such modifications should, of course, only be done *if* there is strong evidence from the congruence of the other characters *and* the details provided by the descriptor that the specimen with the discrepant character really belongs to the same species.

In such cases of failure to match the likelihood of a new species is especially real for the anoles of mainland South and Central America. Both areas are inadequately known, and local new spe-

TABLE 1.

- 1. Head Scales. Smooth: 1,1. Rugose: 2,2. Unicarinate: 3,3. Multicarinate: 4,4. Striate: 5,5.
- 2. Scales between Second Canthals (Fig. 1). 1–30.
- 3. Postrostrals (Figs. 1, 2, and especially 3). 2–15.
- 4. Nasal (Figs. 1, 2, and especially 3). Circumnasal: 0,0. Anterior nasal: 1,1. Divided anterior nasal: 2,2. Inferior nasal 3,3.
- 5. Scales between Nasal and Rostral (Figs. 1, 2, and especially 3). 0-5.
- 6. Scales between Supraorbital Semicircles (Fig. 1). 0-10.
- 7. Enlarged Scales in Supraocular Disk (Fig. 1). 0-30.*
- 8. Elongate superciliaries (Figs. 1, 2, and 4). 0–7.
- 9. Superciliary Series (Figs. 1, 2, and 4). Granules: 1,1. Small scales: 2,2. Larger square or swollen scales: 3,3.
- 10. Loreal Rows (Fig. 2). 1-15.
- 11. Loreal Number (Fig. 2). 2-40.*
- 12. Interparietal Relative to Ear (Figs. 1, 2, and 5). Much smaller: 0,0. Smaller: 1,1. Equal to: 2,2. Larger: 3,3. Much larger: 4,4. Interparietal absent: 5,5.
- 13. Scales between Interparietal and Semicircles (Fig. 1). 0–15. In the absence of an interparietal: 888,888 in the IBM version, NA in the Macintosh version.
- 14. Scales between Interparietal and Nape Scales (Fig. 1). In the absence of an interparietal: 888,888 in the IBM version, NA in the Macintosh version. Count of enlarged scales behind the interparietal distinctly larger than nape scales: 0–15. Scales behind interparietal grading into nape scales: 50,50.*
- 15. Scale Rows between Suboculars and Supralabials (Fig. 2). 0-3.
- 16. Supralabials to below Center of Eye (Fig. 2). 4–15.
- 17. Postmentals (Fig. 6). 1-15.
- 18. Sublabials (Fig. 6). 0-2.
- 19. Sublabials in Contact with Infralabials (Fig. 6). 0–10.
- 20. Dorsals (Figs. 7A, B). Flat, smooth: 1,1. Swollen: 2,2. Unicarinate: 3,3. Multicarinate: 4,4. Triangular or conical crest scales: 5,5.
- 21. Enlarged Middorsal Rows (Figs. 7A, B). 0-30.*
- 22. Middorsal Crests (Fig. 7A). None: 0,0. Low crest: 1,1. High crest: 2,2.
- 23. Flank Scales (Figs. 7A, B). More or less widely separated: 0,0. Juxtaposed: 1,1. Imbricate: 2,2. Heterogeneous: 3,3.
- 24. Size of Ventrals Relative to Dorsals (Figs. 7A, B). Larger: 1,1. Equal: 2,2. Smaller: 3,3. Much smaller: 4,4.
- 25. Smooth/Keeled Ventrals (Figs. 7A, B). Smooth: 1,1. Weakly keeled: 2,2. Strongly keeled: 3,3.
- 26. Ventrals (Figs. 7A, B). Separated: 0,0. Juxtaposed: 1,1. Subimbricate: 2,2. 1mbricate: 3,3.
- 27. Toe Pads (Figs. 8A, B). Pad overlapping first phalanx: 1,1. Pad not distinct from first phalanx: 2,2. No pads: 0,0.
- 28. Lamellar Number (Figs. 8A, B). 0-50.
- 29. Supradigitals. Smooth: 1,1. With indistinct or single keels: 2,2. Multicarinate: 3,3.

TABLE 1. Continued.

- 30. Tail (Fig. 9). Round: 1,1. Weakly compressed: 2,2. Strongly compressed: 3,3.
- 31. Tail Crest (Fig. 9). None: 0,0. Serrate: 1,1. Distinct crest: 2,2. High crest: 3,3.
- 32. Postanals. Present: 1,1. Obscure: 2,2. Absent: 3,3.
- 33. Dewlap (Male) (Fig. 10). Large: 1,1. Intermediate: 2,2. Small: 3,3. Absent: 4,4.
- 34. Dewlap (Female) (Fig. 10). Large: 1,1. Intermediate: 2,2. Small: 3,3. Absent: 4,4.
- 35. Snout-Vent Maximum (Male): 0-300.
- 36. Snout-Vent Maximum (Female): 0-300.
- 37. Tail Length/Body Length: Ratios between 0.8 and 1.2: 1,1. Ratios between 1.3 and 1.7: 2,2. Ratios between 1.8 and 2.2: 3,3. Ratios between 2.3 and 2.7: 4,4. Ratios between 2.8 and 3.0 or more: 5,5.

cies are rather to be expected than not. Only widespread common species are well known, and these only relatively so. This is not true for the anoles of the West Indies, which have been extensively studied, although Cuba, which has been relatively neglected, may be expected to have some efflorescence of novelties.

It is important to realize that the samples for even the best represented species have been selected opportunistically. When readily available, males, females, and juveniles have all been counted and coded, but no attempt has been made to secure a "fair" representation of age and sex classes. Sampling of geographic variation also has been opportunistic: no provision has been made for "adequate" sampling of described subspecies.

It is a matter of course that for poorly known species the samples are always biased and inadequate. There is always, as we have stated in Section II of this paper, extrapolation—judgment—in the association of a second specimen with a unique type or in the association of a population from a new locality with a species previously known from a small, geographically limited range.

It is important to reemphasize that there is no escape from such matters of judgment. Indeed, some of the species in our ANO-LEKEY matrices may be composite. Named subspecies have been consciously lumped in the preparation of the reference matrices for the ANOLEKEY. Some subspecies will, rather surely, be recognized as full, valid species. Others will turn out to be biologically

^{*}The last number is an arbitrary number (see text).

meaningless. We do not attempt to solve such taxonomic problems. We do provide a computer method that will assist the needed comparisons.

The ranges for counts reported in Table 1 as the permissible limits for counts or measurements may deliberately extend beyond the limits actually known for any anole species. Thus, the lowest known count of total number of loreals (character 11, later) is three (in two species of *Phenacosaurus*). I have coded the minimum number for total loreals as two, because this would be a readily expectable variation. Similarly, although the maximum count for lamellae under phalanges ii and iii of the fourth toe (character 28) that I have actually counted is 44, I have coded the maximum for this count as 50. Coding of this kind has been done in several characters in the interest of allowing for easy modification of matrices when known counts are exceeded in either direction.

In the case of character states, Table 1 reports the known states for each character, or, as in the preceding case of certain quantitative characters, extrapolated, to take care of variation that may become known in the future. Any state other than those listed must be coded 888,888 for the IBM version of ANOLE-KEY, or NA (nonapplicable) for the Macintosh version. Any known character state not reported in some individual species matrix but found in some specimen judged to be conspecific with that species may be added to the appropriate matrix by use of the Change Menu in the IBM ANOLEKEY or by simple insertion of the missing coding in the relevant field in the Macintosh *Anolis Handlist*.

Certain counts are more repeatable than others, and some character states may not be readily interpretable. There should be, for example, except for obvious pathological conditions, no equivocal counts for the number of postrostral scales or for the number of scales between the supraorbital semicircles. In other cases, it is very necessary to be aware of arbitrariness and subjectivity in the evaluation of a character. Whenever this is true, advantage should be taken of the program's explicit recognition of the possibility of intermediacy or ambiguity in counts or character states by coding a range even for an individual. In all difficult cases, no time should be wasted by attempting false precision. Code inter-

mediacy, if that seems appropriate, or code the extremes allowed by differing interpretations.

A special case exists in the several instances of graded series of enlarged scales. It is always difficult to decide when to count a scale as enlarged except in the cases of truly abrupt enlargement. In such cases, high counts often reflect a condition much closer to zero than to low counts. I have in such cases used arbitrary high numbers as signals that gradation in size has made counts subjective enough to be meaningless except as indicating gradual change in size.

In the attempt to facilitate the examination of specimens, counts and characters are listed in an order in which a specimen might readily and naturally be examined, beginning with the head and proceeding to the body, limbs, and tail.

Remember that all characters must be recorded by two numbers and that there must be a comma between the numbers. Missing characters should be coded 999,999 in the IBM version, or UA (=unavailable) in the Macintosh version. Nonapplicable characters should be coded 888,888 in the IBM version, or NA (=not applicable) in the Macintosh version. I explicitly reinforce the preceding admonitions by repeating in the coding for character states below double numbers separated by commas.

All terms used in Character Descriptions are defined in Peters' (1964) Dictionary of Herpetology.

Character Descriptions

1. HEAD SCALES. Smooth: 1,1. Rugose: 2,2. Unicarinate: 3,3. Multicarinate: 4,4. Striate: 5,5. Recognition of smooth, more or less parallel ribbed (striate) or rugose head scales or those with strong single ridges (unicarinate) or strong multiple ridges (multicarinate) should not be difficult. The coding for any individual specimen should reflect the most extreme condition in terms of keeling. (The supraoculars may be the most useful scales to look at.)

A very useful character. Striate is the rarest condition.

2. SCALES BETWEEN SECOND CANTHALS (Fig. 1). Range: 1–30. Because they are better defined posteriorly, the canthals are counted from the eye forward. The first canthal always

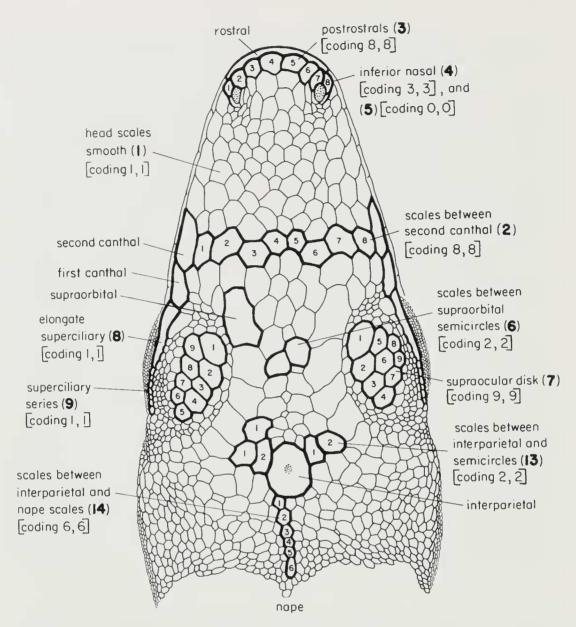


Figure 1. Dorsal view of head of "typical" anole (characters 2–9, 13, and 14). Boldface numbers in parentheses indicate relevant characters; plain numbers in brackets display codings for the illustrated character states.

extends from the canthal ridge backward over the orbit; the second usually does not have such an extension.

This is not necessarily an easy count to make. Variation accounts for some of the problem. At the lower range of counts, variation within a species should not exceed 1 or 2, but it may be 4 or more at the higher extremes. Because two or more scales may be in contact with the second (or the third) canthal on each side, counts within one individual by different observers or by

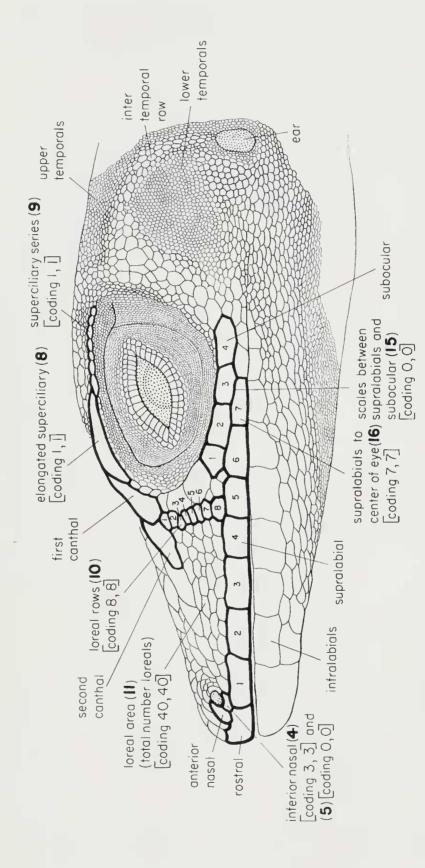
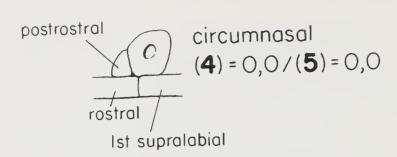
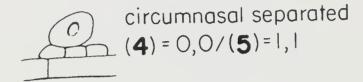
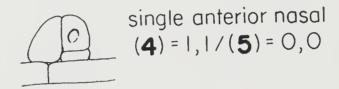
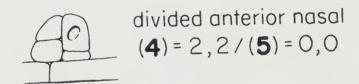


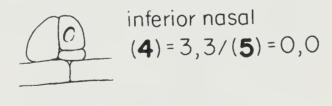
Figure 2. Lateral view of head of "typical" anole (characters 4, 5, 8-11, 15, and 16). Boldface numbers in parentheses indicate relevant characters; plain numbers in brackets display codings for the illustrated character states.











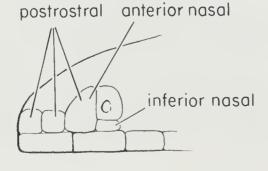


Figure 3. Characters 4 and 5 (nasal scales of anoles).

one observer at different times may differ by 1 or more scales. For matrices based on few specimens, this may not be a useful character; it is primarily useful for the low counts, 2–5, but every species has a characteristic range of counts.

- 3. POSTROSTRALS (Figs. 1, 2, and especially 3). Range: 2–15. Those scales behind the rostral, in contact with it, and between the supralabials are easily counted and will usually not vary within a species by more than 2 to 3. This count may include the circumnasal and any other of the differentiated nasals discussed next.
- 4. NASAL (Figs. 1, 2, and especially 3). Circumnasal: 0,0. Anterior nasal: 1,1. Divided anterior nasal: 2,2. Inferior nasal: 3,3. The nasal in anoles may be a single oval scale that contains the nostril. It is then coded as "circumnasal": 0,0. More often, the scales around the nasal are differentiated, overlap, and at least partially obscure it. Most common is a condition in which a scale anterior to the nasal becomes large and subtriangular and overlaps part of the nasal scale; it is then coded as "anterior nasal": 1,1. In some species, the anterior nasal is divided transversely; it is then reported as a "divided anterior nasal": 2,2. The other scales overlapping the nasal are not considered, except that the inferior nasal—a scale obscuring the lower surface of the nasal—if it comes to overlie the sulcus between the rostral and the first supralabial, is then, and only then, coded "inferior nasal": 3,3. Occasional anomalies (e.g., an anterior nasal replaced by small granules) or conditions due to injury should always be reported as 888,888 in the IBM version, or NA in the Macintosh version.

This and the next (character 5) are important and very useful characters. In some cases both may be difficult to score. It is then best to code the alternatives; these may have been already included in the relevant matrices.

Note that the sulci bounding the scales surrounding and overlapping the nasal are sometimes obscure and that, therefore, the existence, for example, of the anterior nasal must be inferred. Most often, the plausible coding will be 1,1. These difficulties should, in any event, exist only in single specimens or one side of a single individual. Series should obviate or alleviate the problem.

5. SCALES BETWEEN NASAL AND ROSTRAL (Figs. 1, 2, and especially 3). Range: 0-5. In anoline lizards with the condition

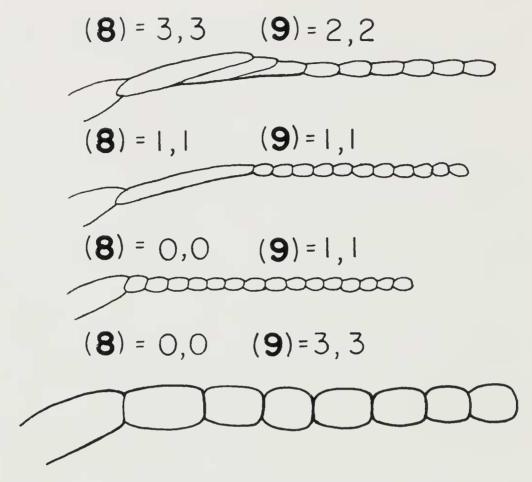


Figure 4. Characters 8 and 9 (superciliary scales in anoles).

"circumnasal"—the simple unmodified nasal scale—that scale may be in contact with the rostral, but one or more scales usually intervene. In the latter case, the minimum number is coded. In the condition "anterior nasal," the nasal itself is obscured, but the anterior nasal is usually in contact with the sulcus between the rostral and first supralabial. This condition is coded 0,0. If the anterior nasal is wholly in contact with the rostral, and the inferior nasal has moved into a position above the sulcus between the rostral and first supralabial, the coding is still 0,0. If, however, a recognizable circumnasal or anterior nasal is separated from the rostral by one or more scales, then the relevant minimum count is recorded.

Often more than one interpretation of any individual condition may be possible. It is, for example, difficult in some species to decide whether the nasal scales are properly regarded as the "anterior nasal" separated by one scale from the rostral (i.e., character 4: 1,1; and character 5: 1,1), or whether the better interpretation is "anterior nasal" divided and in contact with the rostral (i.e., character 4: 2,2; and character 5: 0,0). Ordinarily the ambiguity will already be recorded in the matrix for the species in question, and either of the options will be accepted by the computer.

6. SCALES BETWEEN SUPRAORBITAL SEMICIRCLES (Fig. 1). Range: 0–10. This is a minimum count. Any contact, even a point contact, between the supraorbital semicircles is coded as 0,0.

This is an important count, but in some species there is exceptional variation.

7. ENLARGED SCALES IN SUPRAOCULAR DISK (Fig. 1). Range: 0–30. Enlarged scales arranged as a supraocular disk may be conspicuously larger than any surrounding scales; they are then easily countable. In many species, however, there is gradation such that all counts are subjective; in such counts, a range should be reported. In still other species, the enlargement is so gradual and so limited that there can be no pretense of an accurate count. These cases should be coded arbitrarily as 30,30. If there is no indication of a disk or of significant enlargement of any of the supraocular scales, the coding should then be 0,0.

In many or most species of anoles, this is not the most useful of characters. It is useful in those cases that are unambiguous, i.e., those in which the disk consists of a few large scales that are sharply distinct from surrounding scales. High counts, i.e., smaller scales, are ordinarily subjective.

It is possible to be undecided about whether the coding should be 30,30 or 0,0. Specimens or species in which this occurs should be coded 30,30. The alternative, 0,30, is unacceptable, because the computer will then assume that all counts between 0 and 30 are valid, although no species is known in which so highly variable a condition is true. This coding will be closer to reality.

If you have before you any series containing low counts (large scales; e.g., 3–7) in the supraocular disk and also high counts (small scales; e.g., 20–30), that series should be suspect as composite, and low- and high-count specimens treated separately.

8. ELONGATE SUPERCILIARIES (Figs. 1, 2, and 4). Range: 0-7. The superciliaries are distinguished from the canthals by not

- extending anteriorly beyond the orbit at all. The first superciliary in anoles is usually distinguishable from more posterior superciliaries by being much longer anteroposteriorly. If there is more than one elongate superciliary (there may be as many as six), they will be strongly overlapping and grade in length posteriorly. In a few species, there is no elongate superciliary, and all the superciliary scales are granular. If so the coding is 0,0.
- 9. SUPERCILIARY SERIES (Figs. 1, 2, and 4). Granules: 1,1. Small scales: 2,2. Larger square or swollen scales: 3,3. The scales posterior to the elongate superciliaries—if any are present—may be granular, like the smaller scales of the supraocular area, and are then coded 1,1. Alternatively, there may be one or two rows of distinctly larger but still small scales following the elongate superciliary or superciliaries. If there is difficulty in deciding between these two conditions, or if there is a mixture of granules and small scales, the code should be 1,2. If there is no elongate superciliary and all superciliaries are large and squarish or large and swollen, the conditions should be coded 3,3. The large and squarish and the large and swollen superciliaries are rare conditions, known only in two species in Colombia, Ecuador, and Panama.
- 10. LOREAL ROWS (Fig. 2). Range: 1–15. Loreal rows are counted down from the second canthal or from the first canthal, if it clearly overlaps the loreal area (those scales that continue the subocular arc in front of the eye are preoculars, not loreals, and should never be counted). Sometimes, however, it is difficult to distinguish the preoculars from the loreals; this count is then subjective, by one or two.

This is a count that is often subjective. The next character (total loreal number) is then more useful.

11. LOREAL NUMBER (Fig. 2). Range: 2–40. Total loreal number is easily counted when the loreals are few and the suboculars broadly in contact with the supralabials. Difficulty occurs when the loreals are confluent with a scale row or rows separating the subocular and supralabials. Whenever such a row is complete, the number of loreals will, in most cases, be relatively high (significantly >30); it is then convenient to use the arbitrary coding 40,40. In some cases the preoculars also may be difficult to distinguish. Where the total count would in any case be less than

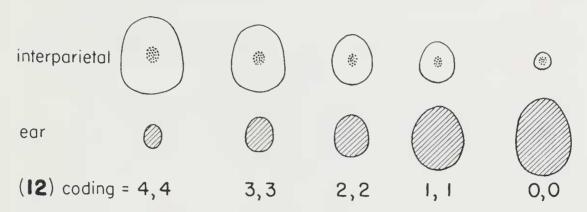


Figure 5. Size of interparietal compared with that of ear (character 12).

40, the possible interpretations may be coded as a range. If the ambiguous situation involves a high count, the arbitrary number 40,40 may be the best solution. Low counts are clearly diagnostic for certain species; even errors of one or two do not matter. High counts do not appear ever to be useful for species discrimination.

12. INTERPARIETAL RELATIVE TO EAR (Figs. 1, 2, and 5). Much smaller: 0,0. Smaller: 1,1. Equal to: 2,2. Larger: 3,3. Much larger: 4,4. Interparietal absent: 5,5. The size of the interparietal relative to the size of the ear is ordinarily unambiguous. If there is ambiguity or variation within a species, code as a range (e.g., 1,3). Again extreme conditions tend to be diagnostic, but in many species codings of 1, 2, or 3 will be appropriate within a single population.

The absence of an interparietal is a rare and apparently derived condition; it is confined to certain species in northwestern South America.

13. SCALES BETWEEN INTERPARIETAL AND SEMI-CIRCLES (Fig. 1). Range: 0–15. In the absence of an interparietal: 888,888 in the IBM version, NA in the Macintosh version. If there is no interparietal, the coding of this character must be 888,888 for the IBM version, or NA in the Macintosh version. This is again a minimum count taken on each side from the interparietal to the nearest scales of the supraorbital semicircles. If the two sides differ, code them as, for example, 2,3. Ranges of 3 or 4 or more are not unusual in certain populations.

14. SCALES BETWEEN INTERPARIETAL AND NAPE SCALES (Fig. 1). In the absence of an interparietal: 888,888 in

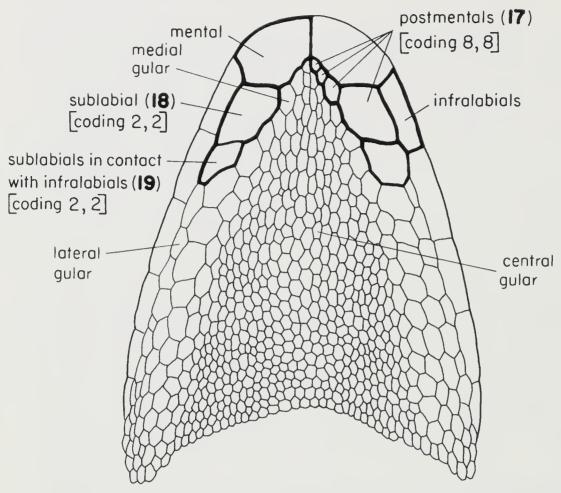


Figure 6. Ventral view of head of "typical" anole (characters 17–19). Boldface numbers in parentheses indicate relevant characters; plain numbers in brackets display codings for the illustrated character states.

the IBM version, NA in the Macintosh version. Count of enlarged scales behind the interparietal distinctly larger than nape scales: 0–15. Scales behind interparietal grading into nape scales: 50,50. This count is made in the approximate midline behind the interparietal, and the scales must be appreciably larger than the nape scales immediately behind them. If there is gradation, and this frequently is the case, the count is surely subjective. If so, code as 50,50.

Discrepancy compared with a matrix is not to be taken seriously in the case of high and subjective counts. However, 0 may have a taxonomic meaning, i.e., an interparietal followed by small scales not significantly different from nape scales. This differs importantly from cases in which scales behind the interparietal

differ abruptly in size from nape scales. Thus, both a low count and an unambiguous count of 0,0 may be very useful species characters.

- 15. SCALE ROWS BETWEEN SUBOCULARS AND SU-PRALABIALS (Fig. 2). Range: 0-3. This is also a minimum count. Any contact between suboculars and supralabials is coded 0,0.
- 16. SUPRALABIALS TO BELOW CENTER OF EYE (Fig. 2). Range: 4–15. Because the posterior termination of the supralabial series may sometimes be difficult to determine, the supralabials are counted from the rostral posteriorly. The sulcus between two supralabials may lie below the center of the eye. The coding is then a range (e.g., 6,7).
- 17. POSTMENTALS (Fig. 6). Range: 1–15. This is a count of all the scales in contact with the mental between the infralabials. It includes the anteriormost sublabials, if these are differentiated.
- 18. SUBLABIALS (Fig. 6). Range: 0–2. A sublabial series is an abruptly enlarged series of scales on each side paralleling or radiating from the infralabials. Abruptly enlarged is here interpreted to require that the first sublabial of each side be at least four to five times the size of the postmental (=medial gular) medial to it. If no such abruptly enlarged scales adjacent to the infralabials exist (i.e., if the postmentals are subequal or grade evenly from larger laterally to smaller medially), or if the enlarged scales are only twice to three times any medial gular, the condition is to be reported as the absence of sublabials, coded 0,0. If both first sublabials are present, the coding is 2,2. If, as sometimes happens, a first sublabial is present on one side only, the coding is 1,1. If there is ambiguity (i.e., if you cannot decide whether or not the putative sublabials are as much as four times the other postmentals), code 0,2.

This character again is a very useful one. If the sublabials are recognized only when the lateral postmentals are four to five times larger than any medial gular, the condition tends to be invariant. Species in which the lateral postmentals are only two to three times larger than the medial gulars tend to be variable in this regard and often have the lateral postmentals subequal to the medial gulars or grading into them.

19. SUBLABIALS IN CONTACT WITH INFRALABIALS (Fig. 6). Range: 0–10. In anoles, the first sublabials, when present,

are almost always in contact with the first infralabials. If the first sublabial is, as rarely may be the case, separated from the first infralabial by a lateral gular, the coding is, of course, 0,0. Any more posterior sublabials (=abruptly enlarged scales in sequence with the first) may also be in contact with the infralabials, or some or all of them may be separated by a row of smaller scales. If only the first sublabials are in contact with infralabials, or there are no posterior sublabials, the coding is 1,1. In some species, there may be as many as 9 in contact, and the number often differs on the two sides (e.g., 1,2 or 6,7). If there are no sublabials at all the coding is, of course, 0,0.

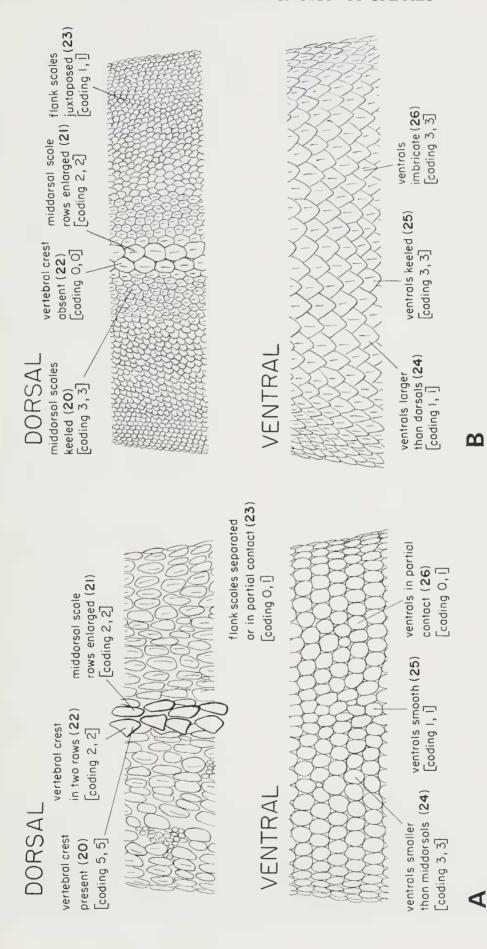
- 20. DORSALS (Figs. 7A, B). Flat, smooth: 1,1. Swollen: 2,2. Unicarinate: 3,3. Multicarinate: 4,4. Triangular or conical crest scales: 5,5. These conditions should be readily recognizable, but if there is any ambiguity, code a range.
- 21. ENLARGED MIDDORSAL ROWS (Figs. 7A, B). Range: 0–30. There may be subjectivity in counts of enlarged middorsal rows. In all such cases, code not single numbers but an appropriate range of values. If the enlarged middorsal rows are not countable because of too gradual a transition to the flank scales, code arbitrarily as 30,30.

As in certain other counts, low counts may be more meaningful and repeatable than high counts.

22. MIDDORSAL CRESTS (Fig. 7A). None: 0,0. Low crest: 1,1. High crest: 2,2. In anoles, "crest" is applied only to one or two rows of sharply enlarged middorsal crests, characteristically triangular or conical, of varying height, but always projecting conspicuously above the paravertebral scales. Most anoles lack such crests.

This is not usually a very useful character. A few anoles have such high crests that I would be remiss if I did not call attention to them.

23. FLANK SCALES (Figs. 7A, B). More or less widely separated: 0,0. Juxtaposed: 1,1. Imbricate: 2,2. Heterogeneous: 3,3. In some, usually giant species, the flank scales are relatively large and may be separated by more or less minute granules. This infrequent condition is coded 0,0. Much more frequently, the flank scales are smaller, and some or all of them are narrowly separated by naked skin or by minute granules, which may allow



Figures 7A, B. Middorsal scales (above) and ventral scales (below) (characters 20-26).

partial contact. This condition is coded 0,1. Many species have the flank scales juxtaposed; this is coded 1,1. If any of the flank scales clearly overlap, the coding is 2,2. In a few species, the larger scales of the flank may be of very unequal size; these are then scored as heterogeneous: 3,3. The minute granules, which may be present in some cases, are not considered in this definition of heterogeneity.

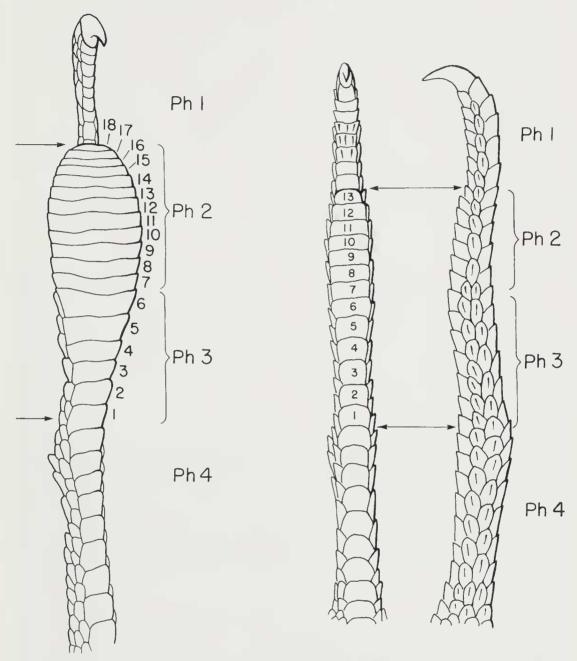
24. SIZE OF VENTRALS RELATIVE TO DORSALS (Figs. 7A, B). Larger: 1,1. Equal: 2,2. Smaller: 3,3. Much smaller: 4,4. Most anole species have the ventrals larger than the largest dorsals. This may, however, be untrue for species with a distinct zone of enlarged dorsals or those species with crest scales. There may even be noticeable variation in this regard within species. As usual, such variation is coded as a range. The condition—ventrals much smaller than dorsals—is known only in Chamaeleolis.

This is often a very useful character.

25. SMOOTH/KEELED VENTRALS (Figs. 7A, B). Smooth: 1,1. Weakly keeled: 2,2. Strongly keeled: 3,3. Some species are truly variable, with ventrals in some individuals keeled (always unicarinate) and in others smooth. Series from such species should be coded 1,2 or 1,3, as appropriate, and individuals either 1,1, 2,2, or 3,3. Other species have keeling so weak that it is a subjective judgment whether or not it exists at all. Such species may be coded 1,2.

There are just a few species in which ventrals vary from smooth to keeled. In most species this is a strong character.

- 26. VENTRALS (Figs. 7A, B). Separated: 0,0. Juxtaposed: 1,1. Subimbricate: 2,2. Imbricate: 3,3. Ventrals are sometimes separated by naked skin or granules and often are juxtaposed or subimbricate (intermediate) or distinctly imbricate (distinctly overlapping). More than one condition can occur within a species or even within an individual. In contrast to the last, this is not a strong character, but some species do have strong separation or strong imbrication.
- 27. TOE PADS (Figs. 8A, B). Pad overlapping first phalanx: 1,1. Pad not distinct from first phalanx: 2,2. No pads: 0,0. In most species of anoles, the adhesive pad under phalanges ii and iii of all the toes projects distally under the proximal infradigital scales of the first (claw-bearing) phalanx. This is coded 1,1. In a number



Figures 8A, B. Two conditions of the digital pads (characters 27 and 28). Phalanges i–iv are labeled, and appropriate counts of lamellae under the pad are illustrated.

of species, not closely related, the projection is minimal or absent, and the proximal termination of the pad is therefore indistinct (coded 2,2). In the adults of one species only (Fig. 8B), a pad is completely lacking (coded 0,0).

28. LAMELLAR NUMBER (Figs. 8A, B). Range: 0-50. Lamellae are the widened distally overlapping scales characteristic

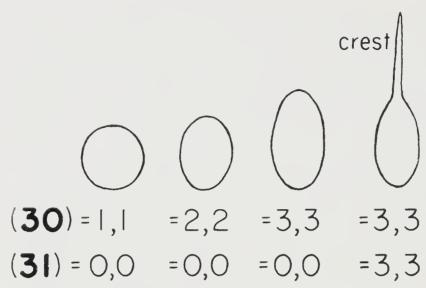


Figure 9. Tail character states (character 30 and 31).

of the adhesive toe pads of anoles. Counts are made only on phalanges ii and iii of the fourth toe of the hind foot. A given range of lamellar counts is species-specific.

Counting lamellar number, as understood here, involves finding the bend between phalanges iii and iv; this is arbitrarily considered the proximal termination of the adhesive pad. A bend also occurs within the pad between phalanges ii and iii, but the distal termination is defined as the joint between phalanges i and ii. Counts are therefore restricted to the lamellae under phalanges ii and iii. At the joint between phalanges iii and iv, the scale at or within the bend is counted. Distally the small terminal lamella of the pad projection is always counted or, in default of such a scale, the scale that lies across the joint between phalanges i and ii.

The intention here is to count only scales of the pad proper. For most anole species, the procedure outlined above probably closely approximates reality. However, scanning electron microscopy has revealed that, in fact, the adhesive lamellae with the functionally adhesive hairs are not necessarily confined within the boundaries of phalanges ii and iii. However, these boundaries provide a convenient macrostructural definition on which to base reproducible counts. Ultimately, however, the justification for this count, rather than counts of all lamellae under the fourth toe, is historical: Boulenger (1885), who first provided standard and recognizable descriptions of anole species, used the count of la-

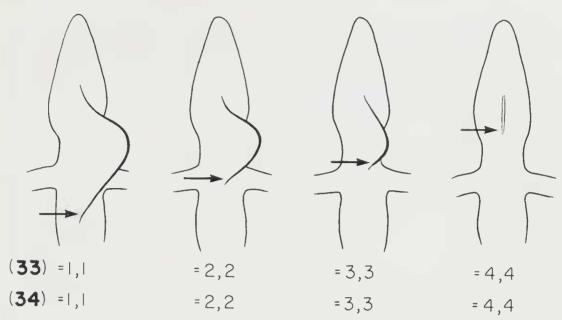


Figure 10. Dewlap extent (characters 33 and 34). Arrow points to posterior termination of dewlap.

mellae under phalanges ii and iii. Stejneger, Cochran, Schwartz, and Williams have routinely used this count.

It is a source of some confusion that workers on Mexican and Central American anoles have rather consistently used other counts, especially the total number of subdigitals under the fourth hind toe. The attempt here is to standardize lamellar counts with those conventionally used for West Indian and South American anoles.

- 29. SUPRADIGITALS. Smooth: 1,1. With indistinct or single keels: 2,2. Multicarinate: 3,3. Most anoles have multicarinate supradigital scales. The alternative conditions—smooth (1,1) or with indistinct, usually single keels (2,2)—are relatively uncommon and, therefore, when they do occur, diagnostic.
- 30. TAIL (Fig. 9). Round: 1,1. Weakly compressed: 2,2. Strongly compressed: 3,3. Strongly compressed tails are readily recognized. Weakly compressed or round tails are common.
- 31. TAIL CREST (Fig. 9). None: 0,0. Serrate: 1,1. Distinct crest: 2,2. High crest: 3,3. Round tails have no crests. Weakly compressed tails may or may not be serrate above. Strongly compressed tails may or may not have a crest, and may not even be serrate above, and, if present, the crest may be high or low. Fe-

males rarely have high crests and variation from population to population within a species and within a sex is not unusual.

- 32. POSTANALS. Present: 1,1. Obscure: 2,2. Absent: 3,3. The characteristic enlarged postanal scales of male anoles are usually laterally expanded ovals, somewhat resembling airplane propellers, a variable distance behind the vent. They may also be round, large or small, obscure, or, as in males of not a few species, and always in females, absent.
- 33. DEWLAP (Male) (Fig. 10). Large: 1,1. Intermediate: 2,2. Small: 3,3. Absent: 4,4. Dewlaps should be scored as large if they extend onto the belly past the insertion of the arms, intermediate if they reach just the level of the axilla, and small if they are shorter than this. If the animal is female, code 888,888 in the IBM version, or NA in the Macintosh version.
- Large (1,1), as defined here, is a very inclusive term, and species so coded will differ conspicuously in the dewlap area. The present definition has been adopted to avoid the problem of quantifying dewlap area and to permit the scoring of dewlaps that are not extended in preserved animals.
- 34. DEWLAP (Female) (Fig. 10). Large: 1,1. Intermediate: 2,2. Small: 3,3. Absent: 4,4. Females may possess a dewlap, which is then scored just as in males. However, both an indication—a mere zone of differentiated scales—or a mere fold (i.e., any structure that could not be substantially extended) and a complete absence should be scored 4,4. If the animal examined is a male, code 888,888 in the IBM version, or NA in the Macintosh version.
- 35. SNOUT-VENT MAXIMUM (Male). Range: 0-300. Snoutvent length should be measured on every specimen. The number with which this length is compared in the coded matrix is the maximum known for the relevant sex in each species. Only if the matrix number is exceeded by a substantial margin should a tentative identification be rejected.

A 0 is used as the minimum for size only because hatchling size is unknown for most anoles.

A female should be coded 888,888 in the IBM version, or NA in the Macintosh version.

36. SNOUT-VENT MAXIMUM (Female). Range: 0-300. The principles suggested for the interpretation of the comparable entry in males apply here also. For a male, code this entry 888,888 in

the IBM version, or NA in the Macintosh version. As in males, 0 is treated as the minimum only because hatchling size is nearly unknown or unreported in anoles.

37. TAIL LENGTH/BODY LENGTH. Ratios between 0.8 and 1.2: 1,1. Ratios between 1.3 and 1.7: 2,2. Ratios between 1.8 and 2.2: 3,3. Ratios between 2.3 and 2.7: 4,4. Ratios between 2.8 and 3 or more: 5,5. Tail length should be measured whenever the tail is neither broken nor regenerated. (The regenerated portion of a lizard tail always differs in scale form from the unregenerated portion; such difference in scale form is never seen in an unregenerated tail.)

The measured tail length must then be compared with the snoutvent length, and the ratio of tail length over body (snout-vent) length obtained. Ratios between 0.8 and 1.2 should be coded 1,1. Ratios between 1.3 and 1.7 may be coded 2,2; those between 1.8 and 2.2 coded 3,3; those between 2.3 and 2.7 coded 4,4; and those between 2.8 and 3 or more 5,5. All intermediates should be so coded (e.g., individuals with ratios 2.75 as 4.5; series with ratios ranging from 1.5 to 1.9 as 2,3).

V. ANOLEKEY-OPERATIONAL INSTRUCTIONS

ERNEST E. WILLIAMS

The ANOLEKEY is menu-driven and is available in two versions, one in Basic for the IBM PC, written by Hugh and Stanley Rand, and one in HyperCard for the Macintosh, written by Robert O'Hara. Both versions do essentially the same job but differ in one aspect. The HyperCard version for the Macintosh is essentially a database for all anoles with an ANOLEKEY as one of the options and the reference matrices as part of the database.

At this time, most reference matrices are not finished and lack the added feature of descriptors. Only Puerto Rico, Jamaica, Ecuador, and Costa Rica-Panama subsets of the total database are available at this time. However, it is hoped to add other subsets in the near future.

Because the ANOLEKEY itself is functional, and we welcome suggestions and improvements, we have decided to go public at this time. A copy of ANOLEKEY and associated documentation (Hugh and Stanley Rand, 1994, for the IBM PC, and Robert

O'Hara and Ernest Williams, 1994, for the Macintosh) is available to anyone interested in trying the ANOLEKEY and willing to report bugs and/or suggest improvements. Each copy will include a description of the 37 characters and how to count them (Williams, this paper). (For the IBM version, write to A. S. Rand [include formatted disk]; for the Macintosh version, write to E. E. Williams [include formatted disk].)

In addition to the ANOLEKEY and its documentation, the distribution disk for the IBM includes a version of the Basic Program that is not customized for use with the anoles but that can be used with any group. For a comparable bare-bones HyperCard for the Macintosh version, write to E. E. Williams.

VI. ANOLEKEY IN BASIC FOR THE IBM PC

A. STANLEY RAND

The IBM version of ANOLEKEY consists of a Basic Program, a "Readme" file that describes the use of the program, a text file describing the 37 anole characters, and several reference matrices, all stored on a single floppy disk. The program was created in MS-DOS, is stored in ASCII, and may be loaded and run using BASICA, Turbo BASIC, or Quick Basic. The program should run on any IBM PC or clone.

The disk contains two versions of the program adapted to identify anoles, one that prints out a hard copy of results and the other that does not. There is also a stripped down version that lacks the modifications for use with anoles so that it can be used for any group.

The reference matrices each contain information on all the recognized species of anoles from a specific geographic area. Four areas are currently available: Puerto Rico, Jamaica, Panama—Costa Rica, and Ecuador. Each matrix contains the ranges for the 37 standard characters for each species and the short list of diagnostic characteristics, the "descriptor," for each species.

The program allows you to load a reference matrix (e.g., Pan-ama-Costa Rica) and then enter the information for the 37 standard characters for an individual or series to be identified. Most of these characters are scale characters, two involve size, and one involves proportions. A list of these characters and their codings is given in Appendix 2 of the "Readme" file. Some of these

characters are counts that are entered directly; the others must be coded. You are prompted for each character and informed of the possible codings. For each character, the program will accept only an appropriate range of values. You must enter for each character two numbers separated by a comma, a minimum and a maximum. This allows you to enter information from a series of animals, to code the two sides of an animal if they are different, and to code an ambiguous situation. The data for the unknown can be corrected and stored in a file.

Once the data are entered and checked, the program asks you to enter the number of characters where a species in the reference matrix with which the unknown is being compared must differ before being eliminated as a possible match. The larger this number, the more species are matched and the more different they may be from the unknown.

The program then compares the unknown with the reference matrix and produces a report that lists the species that match, giving for each the total number of characters in which it differs, what the differing characters are, and how great the differences, positive or negative, are. For each matching species the program supplies a short description of characters such as color, distribution, and any peculiarities that could help in identification.

The program allows you to create or change a reference matrix, although creating a reference matrix is probably more easily done with a word-processing program outside of ANOLEKEY. The details of the format for a reference matrix are given in the "Readme" file.

As a shortcut, there is also the option of comparing an unknown with the reference matrix using only two characters, instead of all 37. This will be useful if the unknown has among its 37 standard characters one or two states that are so rare or unusual that only one, two, or at most three species can possibly be matches. The descriptor or descriptors will confirm the identification.

VII. THE ANOLIS HANDLIST: A HYPERCARD VERSION OF THE ANOLEKEY

ROBERT J. O'HARA AND ERNEST E. WILLIAMS

The Macintosh version of the ANOLEKEY is much more than a computer key, and to better indicate its multiple functions we

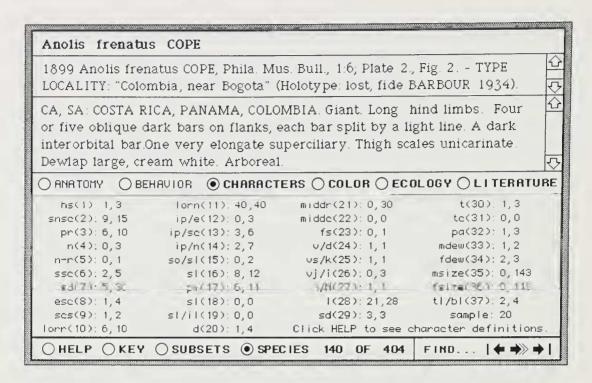


Figure 11. Species card for *Anolis frenatus* from the *Anolis Handlist*. The buttons at the lower left, present on every card in the stack, allow the user to choose the Help, Key, Subset Editor, or Species cards. The diagnostic characters are now being shown on this card; information on anatomy, behavior, color, ecology, or literature may be displayed by clicking on the appropriate button.

refer to it as the *Anolis Handlist*. In addition to serving as a key, the *Anolis Handlist*, when completed, is intended to be an index and checklist for the possibly more than 400 species of *Anolis*, *Chamaeleolis*, *Chamaelinorops*, and *Phenacosaurus* residing in North, Central, and South America.

Although the *Anolis Handlist* is still incomplete, as already mentioned, we describe it here as though it were complete. While waiting for EEW to complete additional reference matrices along with the several subsidiary COLOR, etc., fields, described below, the ANOLEKEY is fully functional.

The Anolis Handlist was created with Apple Computer's HyperCard, an object-oriented software development environment that makes use of all of the standard elements of the Macintosh interface, including pull-down menus, windows, buttons, and scroll bars (Apple Computer, 1987). HyperCard documents may be thought of as groups of interactive electronic file cards:

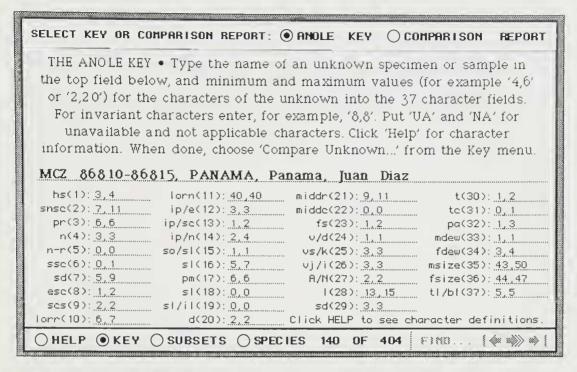


Figure 12. The Key card, with data from a series to be compared entered into the character fields.

each HyperCard document is known as a "stack," and each stack is made up of one or more individual screens, or "cards" (Apple Computer, 1989). The *Anolis Handlist* is a single HyperCard stack, and it consists of more than 400 Species cards (some of the species are undescribed or not well understood), a Key card, a Subset Editor card for use with the Key, and a Help card that provides general information as well as definitions of diagnostic characters (see Section V, earlier).

Each Species card displays in the upper two fields the name, author, original citation, and type locality and in the bottom two fields two kinds of characters: (1) the "descriptors," those that are especially diagnostic, i.e., special for individual species, and (2) the characters that are routinely recorded for all species (Fig. 11). Also on each Species card, provision is made for information on anatomy, behavior, color, ecology, and literature, and this information may be recorded and read on fields called up by choosing the corresponding button the middle row between the upper and lower fields. The Species cards may be browsed one by one using the navigational buttons in the lower right corner

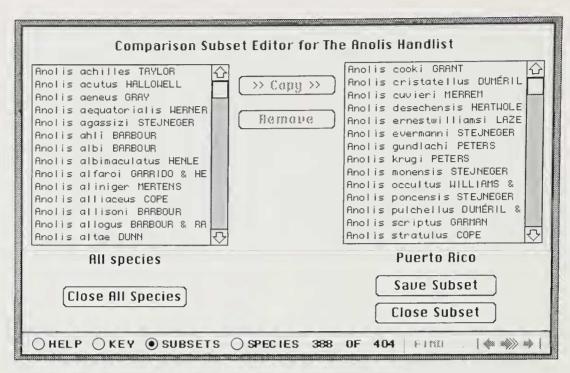


Figure 13. The Subset Editor card. The scrolling field on the left lists all the species in the *Anolis Handlist*, and the field on the right contains one of the comparison subsets. These subsets may be edited and saved for future use with the Key.

of the cards, and they may be sorted according to species name, year of description, or any chosen diagnostic character by making the appropriate selection from a pull-down menu. Species cards may be added, deleted, or edited at the discretion of the user. Every scrolling field on each Species card can hold up to 30,000 text characters; thus, the field for original citation could in fact contain a full synonymy, or the literature field an extensive bibliography. All the data in the Species cards may be searched at will by selecting the FIND button (Fig. 11).

The Key card (Fig. 12) permits the user to enter values for the diagnostic characters of an unknown specimen or series and to compare this unknown to all the Species cards or to a subset of the Species cards. When going to the Key card from a Species card, the option is provided to automatically read the data from that Species card into the Key as if it were an unknown, thus permitting comparisons among known species as well as among unknowns and knowns. Once the characters of a specimen or

species to be compared have been entered on the Key card, the user selects "Compare Unknown" from a pull-down menu and responds to a series of prompts asking, for example, by how much the unknown may differ from the reference data on the Species cards and still be considered a match.

The Subset Editor (Fig. 13) is an adjunct to the Key card. Comparison of an unknown to all of the Species cards in the

A report from The Anolis Handlist by E. E. Williams Sunday, April 10, 1994, 4:48 PM

Anolis auratus DAUDIN

CA, SA: eastern and central PANAMA, COLOMBIA, ECUADOR, VENEZUELA, BRAZIL. Small. Toepads not overlapping first phalanx. Light lateral line. Dewlap large, blue or black. Sharply enlarged keeled middorsals. In grass.

''MCZ 86810-86815, PANAMA, Panama, Juna Diaz'' misses by 0 out of 37 key characters.

• Anolis biporcatus (WIEGMANN)

CA, SA: MEXICO, GUATEMALA, EL SALVADOR, NICARAGUA, COSTA RICA, PANAMA, COLOMBIA. Large. Heavy bodied. Green changing to dark brown. Dewlap moderate in male, white basally, mostly powder blue with red-orange free margin, in female smaller, sometimes with black flecks. In canopy.

''MCZ 86810-86815, PANAMA, Panama, Juan Diaz'' misses by 5 out of 37 key characters:

n(4) by 1, n-r(5) by -1, A/N(27) by 1, 1(28) by -1, t1/b1(37) by 1

• Anolis humilis PETERS

CA: NICARAGUA, COSTARICA, PANAMA. Small, Brown. Dewlaplarge, red with bright yellow margin. Axillary pits. About 10 middorsal rows enlarged, larger than ventrals, the two median rows smallest. On or near ground.

''MCZ 86810-86815, PANAMA, Panama, Juan Diaz'' misses by 5 out of 37 key characters:

lorn(11) by 1, d(20) by -1, v/d(24) by -2, A/N(27) by 1, t1/b1(37) by 3

Anolis intermedius PETERS

CA: COSTARICA, PANAMA. Small, Brown or greyish. Dewlap bone white. Arboreal.

^{&#}x27;'MCZ 86810-86815, PANAMA, Panama, Juan Diaz'' compared with the subset ''PANAMA.''

''MCZ 86810-86815, PANAMA, Panama, Juan Diaz'' misses by 4 out of 37 key characters:

lorn(11) by 13, so/sl(15) by 1, A/N(27) by 1, tl/bl(37) by 2

• Anolis kemptoni DUNN

CA: PANAMA. Small. Greyish brown. White line under the eye. Dewlap skin red with orange anterior spot, scales whitish. Arboreal.

''MCZ 86810-86815, PANAMA, Panama, Juan Diaz'' misses by 3 out of 37 key characters:

so/sl(15) by 1, A/N(27) by 1, tl/bl(37) by 2

• Anolis lemurinus COPE

CA: MEXICO, GUATEMALA, HONDURAS, NICARAGUA, COSTA RICA, PANAMA. Moderate size. Olive brown with dark dorsal blotches or (females diamond-shaped middorsal markings or a black-edged middorsal stripe. Dewlap dark red with black scales. In Panama only in western and central regions.

''MCZ 86810-86815, PANAMA, Panama, Juan Diaz'' misses by 5 out of 37 key characters:

n(4) by 1, d(20) by -1, A/N(27) by 1, 1(28) by -1, t1/b1(37) by 2

• Anolis lionotus COPE

CA: PANAMA. Moderate size. Light lateral line. A dorsal zone of about 10 rows of enlarged smooth scales much larger than ventrals. Dewlap large, orange. Semiaquatic. Only central Panama.

''MCZ 86810-86815, PANAMA, Panama, Juan Diaz'' misses by 5 out of 37 key characters:

n(4) by 1, d(20) by 1, v/d(24) by -2, A/N(27) by 1, t1/b1(37) by 2

• Anolis poecilopus COPE

CA, SA: PANAMA, COLOMBIA. Moderate size. Light lateral line. Dewlap large, orange. Head scales small. A dorsal zone of about 20 rows of enlarged keeled scales about as large as ventrals. Semiaquatic. In Panama only in eastern region, in Colombia only in western region.

''MCZ 86810-86815, PANAMA, Panama, Juan Diaz'' misses by 4 out of 37 key characters:

snsc(2) by -3, n(4) by 1, A/N(27) by 1, t1/b1(37) by 2

• Anolis tropidogaster HALLOWELL

CA, SA: PANAMA, COLOMBIA. Small. Often an indication of a light middorsal stripe in both sexes. Dewlap yellow, orange or reddish. Trees and bushes.

 $\mbox{''MCZ~86810-86815},\mbox{ PANAMA, Panama, Juan Diaz'' misses by 5 out of 37 key characters:$

n(4) by 2, ssc(6) by -1, d(20) by -1, A/N(27) by 1, t1/b1(37) by 2

• Anolis vittigerus COPE

CA, SA: PANAMA, COLOMBIA. Moderate size. Variable and complex pattern on nape. Dewlap with central dark spot. In Panama only in eastern region.

''MCZ 86810-86815, PANAMA, Panama, Juan Diaz'' misses by 5 out of 37 key characters:

middr(21) by -1, A/N(27) by 1, 1(28) by -1, mdew(33) by -1, t1/b1(37) by 1

• Anolis woodi DUNN

CA: COSTA RICA, PANAMA, Cordillera Talamanca. Large. Olive with indistinct rusty spots. Dewlap pink orange at edge, amber yellow in middle, bluish white at base.

''MCZ 86810-86815, PANAMA, Panama, Juan Diaz'' misses by 4 out of 37 key characters:

n(4) by 1, ip/e(12) by 2, A/N(27) by 1, t1/b1(37) by 1

Figure 14. A Comparison Report from the *Anolis Handlist*. Reports may be edited and printed with the standard HyperCard facilities or copied into any word processor for editing and printing.

Anolis Handlist may be time-consuming, and in most cases some additional information about the unknown (such as its collecting locality) or its species group will permit a comparison to be restricted to some subset of the total collection of species. The Subset Editor allows the user to assemble comparison subsets, either manually or automatically, to save these subsets for future use and to edit them as necessary. When making a comparison with the Key, the user may specify that the comparison be made against one of the existing subsets.

The result of a comparison made with the Key is a Comparison Report (Fig. 14). This report specifies the name of the unknown that was compared, the name of the subset (if any) to which it was compared, the number of characters that differ between the unknowns and the knowns, and by how much these characters differ. The Comparison Report may be printed using the PRINT

FIELD option under FILE. (PRINT CARD will print only the portion of any field visible initially.)

The *Anolis Handlist* has been designed specifically for the Macintosh computer. The principles it embodies are general ones, however, and these principles could be implemented on a variety of computer platforms.

ACKNOWLEDGMENTS

We acknowledge with special gratitude the struggles of two anonymous reviewers to catch errors, typographical and other, revise the logic of our discussions, and, with all good will, to generally emend and improve our paper. Nancy Knowlton and John Cadle saw a much later version. We also greatly appreciate their helpful comments.

LITERATURE CITED

- APPLE COMPUTER. 1987. Human Interface Guidelines: The Apple Desktop Interface. Reading, Massachusetts, Addison-Wesley Publishing Company.
- ——. 1989. HyperCard Stack Design Guidelines. Reading, Massachusetts, Addison-Wesley Publishing Company.
- BOULENGER, G. A. 1885. Catalogue of the lizards in the British Museum (Natural History). London, Trustees of the British Museum, 2nd ed. 2: xiii + 497 pp.
- Knowlton, N. 1993. Sibling species in the sea. Annual Review of Ecology and Systematics, 24: 189–216.
- LAZELL, J. D., JR. 1983. Biogeography of the herpetofauna of the British Virgin Islands, with description of a new anole (Sauria: Iguanidae), pp. 99–117. *In* A. G. J. Rhodin and K. Miyata (eds.), Advances in Herpetology and Evolutionary Biology. Cambridge, Museum of Comparative Zoology.
- Morse, L. E., J. A. Peters, and P. B. Hamel. 1971. A general data format for summarizing taxonomic information. BioScience, 21: 174–181.
- Peters, J. A. 1964. Dictionary of Herpetology. New York, Hafner Publishing Company. ix + 392 pp.
- Peters, J. A., and B. B. Collette. 1968. The role of time-share computing in museum research. Curator, 11: 65–75.
- Peters, J. A., and R. Donoso-Barros. 1970. Catalogue of the Neotropical Squamata: Part II. Lizards and amphisbaenians. United States National Museum Bulletin, 297: 1–293.
- Peters, J. A., and B. Orejas-Miranda. 1970. United States Museum, Part I. Snakes. United States National Museum Bulletin, 297: 1–347.
- RAND, A. S., AND E. E. WILLIAMS. 1969. The anoles of La Palma: Aspects of their ecological relationships. Breviora, Museum of Comparative Zoology, 327: 1–19.

- WILLIAMS, E. E. 1972. The origin of faunas. Evolution of lizard congeners in a complex island fauna: A trial analysis. Evolutionary Biology, 6: 47–89.
- ——. 1976. South American anoles: The species groups. Papeis Avulsos de Zoologia, São Paulo, **29**(26): 259–268.
- ——. 1983. Ecomorphs, faunas, island size and diverse end points in island radiations of *Anolis*, pp. 326–370. *In* R. B. Huey, E. R. Pianka, and T. W. Schoener (eds.), Lizard ecology, studies of a model organism. Cambridge, Harvard University Press.